Society as a moral order: Adam Smith's theory of sociability as a response to Mandeville and Rousseau
Ivan Prates Sternick

https://doi.org/10.1590/0103-6351/7560

Submitted on: 2023-06-29
Posted on: 2023-06-30 (version 1)
(YYYY-MM-DD)
Society as a moral order: Adam Smith's theory of sociability as a response to Mandeville and Rousseau

Abstract

This paper aims to show that Smith’s theory of sympathy and the impartial spectator can be understood as an attempt to overcome the selfish anthropology advocated by Mandeville and denounced by Rousseau. In Smith’s view, if Mandeville’s theory of the psychology behind commerce and exchange was correct, then Rousseau would be right in his denunciation of the moral evils of civilization. However, for Smith, Mandeville’s theses were wrong, and thus Rousseau’s critiques were largely unfounded, because, quite paradoxically, they relied on Mandeville’s description of sociability. Therefore, the often emphasized sympathies of Smith for Rousseau’s arguments should be mitigated.

Keywords: Adam Smith, Sociability, Self-love, Jean-Jacques Rousseau, Bernard Mandeville.

JEL Codes: B10, B11, B12.

Resumo

Este artigo busca mostrar que a teoria da simpatia e do espectador imparcial de Smith pode ser entendida como uma tentativa de superar a antropologia egoísta defendida por Mandeville e denunciada por Rousseau. Na visão de Smith, se a teoria de Mandeville a respeito da psicologia por trás do comércio e da troca estivesse correta, então Rousseau teria razão em sua denúncia dos males morais da civilização. No entanto, para Smith, as teses de Mandeville estavam equivocadas, e, assim, as críticas de Rousseau eram amplamente infundadas, pois, paradoxalmente, elas se baseavam na descrição da sociabilidade apresentada por Mandeville. Nesse sentido, as frequentemente enfatizadas simpatias de Smith pelos argumentos de Rousseau devem ser mitigadas.

Palavras-chave: Adam Smith; Sociabilidade; Amor-próprio; Jean-Jacques Rousseau; Bernard Mandeville.

Códigos JEL: B10, B11, B12.
1 Introduction

In *The Theory of Moral Sentiments* (TMS), Adam Smith joins a debate about the nature of human sociability, establishing a critical dialogue with different thinkers of the 17th and 18th centuries, among them Mandeville and Rousseau. In this debate, sociability is understood as the qualities and attributes of human nature that dispose individuals to live together, to relate peacefully to one another and to promote mutual good offices. The problem of sociability also takes the form of a reflection on the foundation of moral distinctions, social obligations and society itself.

It is important to note that, in this period, the theories of sociability played a relevant role in the formation of political economy as an autonomous object of theoretical reflection. Ultimately, what is at stake in the discussion of sociability is the possibility of thinking society, cooperation, commerce, and the accumulation of capital as logically prior to and analytically independent from the impositions of political power. If there is some principle in human nature that naturally inclines individuals to peaceful cooperation; if “spontaneous” human interaction has an order and an ethical principle of its own; then it is possible to represent “economic” interaction as a relatively autonomous phenomenon vis-à-vis the state. Otherwise, this interaction needs to be ordered by impositions of a superior power, be it political or divine. In this case, the economic phenomenon loses its relative autonomy. In this sense, the discussions about a pre-government sociability gave rise to the possibility of political economy as a relatively independent object of science.

In the case of Smith’s moral philosophy, the decisive issue is to think about the ethical foundations of self-love and commerce. A large number of interpreters stress that this is done mainly on the basis of a critique of Mandeville’s selfish theses on human nature and sociability (Colletti, 1972; Hundert, 1994; Winch, 1996; Force, 2003; Hurtado, 2004; Cerqueira, 2008; Hont, 2015; Douglass, 2017; Griswold, 2018; McHugh, 2018; Bee, 2021), although for some, Smith’s theory of commercial society would ultimately be too close to Mandevillian premises (Hirschman, 1977; Dupuy, 1987; Hundert, 1994; Kerkhof, 1995; Hurtado, 2004; Douglass, 2017). Among all these, there are those who also recognize the importance of Rousseau for Smith’s theoretical development, with some ascribing to them completely antagonistic views of man and society (Colletti, 1972; Ignatieff, 1984; Hurtado, 2004; McHugh, 2018). Whereas others seek to bring them
closer, to varying degrees, highlighting the possible influence of Rousseau’s diagnosis of the evils of inequality and self-love on Smith’s thinking, despite their major divergences (Force, 2003; Rasmussen, 2006; 2008; Hanley, 2008; 2009; Hont, 2015; Stimson, 2015; Griswold, 2018). Sagar (2018a), in turn, holds that Rousseau’s ideas were not an important influence to Smith’s theoretical development and did not pose a serious challenge to his thought, since they were neither novel nor sophisticated in comparison to Mandeville’s and Hume’s moral theories.

This paper aims to show that Smith’s theory of sympathy and the impartial spectator was developed as a response to the problem of the moral status of the sociability founded on self-love, as bequeathed by the selfish anthropology advocated by Mandeville and denounced by Rousseau. In Smith’s view, if Mandeville’s theory of the psychology behind commerce and exchange was correct, then Rousseau would be right in his eloquent denunciation of the moral and political evils of civilization. However, for Smith, Rousseau was wrong because, quite paradoxically, he adhered too closely to Mandeville’s theses, which he wanted to criticise.

Thus, in contrast to those who see complete opposition between Smith and Rousseau, this means that the former agrees with the latter’s criticism of Mandeville’s anthropology and his idea of sociability as the desire to obtain external advantages and the esteem of others at any cost. Rousseau’s critique can be considered important for Smith – contrary to Sagar’s (2018a) statement – inasmuch as it made explicit the moral and political consequences of Mandeville’s conception of sociability. Nevertheless, in Smith’s view, Rousseau was wrong to incorporate a Mandevillian-inspired psychology into his description of the workings of modern society, since it does not correctly explain human sociability and interaction. Therefore, Smith’s sympathies for Rousseau’s critiques of commercial society should be qualified.

The first section analyzes Smith’s comments on Mandeville and Rousseau in the *Letter to the Authors of the Edinburgh Review* (1756). The aim is to show how, in Smith’s reading, Mandeville’s anthropology strips sociability of an authentic ethical content, serving as the basis for a pernicious representation of society, and how Rousseau draws on this premise to craft a “sublime” critique of civilization. The second section shows how Smith elaborates his theory on the moral foundations of sociability from a critical engagement with Rousseau’s critiques of the Mandevillian conception of men, comparing

---

1 See also Paganelli, Rasmussen, and Smith (2018), which contains a collection of articles on Smith and Rousseau.
the respective notions of pity and sympathy. The third section aims at explaining how this
theory unfolds into a theory of conscience, and how the latter is articulated with an idea
of society as an order ruled by moral laws that emerge out of the sociability based on self-
love as the desire for deserved esteem. The fourth section analyzes how the notion of the
impartial spectator can be understood as a response to the problem posed by Mandeville’s
understanding of the role of the desire for esteem as a vehicle of socialization, and to its
consequences for human personality and morality, as denounced by Rousseau.

2 Mandeville and Rousseau in the Letter to the Authors of the
Edinburgh Review (1756)

In 1756, Smith sent an anonymous letter to his friend Alexander Wedderburn, then
editor of the newborn *Edinburgh Review*. His general aim was to suggest to the editors
that they broaden the scope of the journal, which originally contemplated critical reviews
of Scottish and English works published in the previous semester, so that they would also
consider works published on the continent, and include only Scottish works that were
“tolerably decent”. Smith suggests that, by considering relevant works from the
continent, the editors would contribute to raising Scotland’s reputation among nations
that cultivate the letters and sciences. This task would not be very laborious, for it was
mainly in England and France that something of originality and relevance had hitherto
been produced in literature, natural philosophy, and moral philosophy (EPS, 242-3). To
illustrate his point, he sets out to analyze and compare some major recent contributions
from those two countries.

At the end of the *Letter*, in reviewing the major French contributions in the field
of moral philosophy, Smith stresses the importance of Rousseau’s recently published
*Discourse upon the origin and foundation of the inequality amongst mankind* (1755;
*Second Discourse*) (EPS, 249-250). Smith claims that Rousseau’s work was inspired by
the second volume of Mandeville’s *Fable of the Bees*, published in 1729, except that in
the *Second Discourse*, “the principles of the English author are softened, improved, and
embellished, and stript of all that tendency to corruption and licentiousness which has

---

2 This document is part of the third volume of the *Glasgow Edition* of Smith’s works (*A Letter to the Authors
of the Edinburgh Review*, p. 242-256), entitled *Essays on Philosophical Subjects* (Smith, 1980; hereafter
EPS). References to it and to Smith’s other works will follow the pattern established by the *Glasgow
Edition*. In the case of EPS, references will be made as follows: the abbreviation of the work followed by a
comma and the page number in question (e.g., EPS, 251).

3 On the general content of the letter, see Lomonaco (2002).
disgraced them in their original author” (EPS, 250). He then goes on to compare the ideas of the two authors.

It is above all the problem of sociability that is at stake in the parallels that Smith draws between Rousseau and Mandeville, as well as in the three passages from the *Second Discourse* that were translated and inserted at the end of his commentary (cf. EPS, 251-4). Smith was interested in the way in which both conceived human nature and elaborated a natural history of sociability and society (see also Griswold, 2018, pp. 97). We will try to show that his commentary on the relationship between the two is very illuminating of the movement he makes in TMS in developing the moral foundations of sociability.

First of all, it is important to underline the central issue that is at stake in the second volume of *The Fable of the Bees*. That book, mentioned by Smith as the inspiration for Rousseau’s *Second Discourse*, is precisely the one in which Mandeville develops more deeply the views on the foundations of sociability he had presented in the first volume. As stressed by Smith, Mandeville (1988, ii.177, 180-5) states that human beings do not have a natural inclination to live in society. According to him, their desire to live together is based on the concern for their own happiness, or on the desire to better their own condition, in view of their helplessness to satisfy their own needs alone. Therefore, in his view, what makes human beings sociable is their concern for themselves, and not for others⁴.

According to Mandeville, the foundation of society is human needs, or natural and moral want (*ibid.*, i.344-5, ii.348-350). Self-love, understood primarily as the desire for moral superiority, alongside material needs, is the vehicle of sociality. Pride, or desire for esteem at any cost, is the passion mobilized in the process in which individuals are taught to bend their selfish affections, to obey social rules, and to practice good offices (*ibid.*, ii.65, 74-5). Socialization involves the concealment of selfish motivation behind a “specious cloak of sociableness”, the artificial veil of civility and politeness, through which individuals relate to one another. In other words, the process of becoming sociable entails a split between being and appearing, which makes individuals closed and hypocritical beings (*ibid.*, i. 234-5, 349). This means that sociability is merely apparent: individuals dissimulate good behavior to gain others’ esteem, but are often willing to deceive them for their own benefit when they can do so without getting caught.

---

⁴ It should be remarked that Mandeville gets into a long consideration about the conjectural history of society and sociability, which will not be considered here since it does not bear directly on the main point of the article.
Therefore, according to Mandeville, social intercourse, or human “commerce” (in the broad sense), is emptied of any authentic moral content. This is not only because it is driven by vices - such as avarice, envy, and pride - and because virtues are only apparent, or disguised in view of a selfish end. But in the sense that, ultimately, human motivation is not inspired by properly moral principles, but by the desire to be esteemed superior at any cost. The practice of virtue, of what is right and just, is not an end in itself, but a means to the attainment of pleasure.

From this selfish foundation, Mandeville arrives at the division of labor and exchange. As soon as property is secured by a recognized authority, men, guided by the desire to improve their condition, begin to divide labor and exchange their products (ibid., ii.284). Society is constituted by the services that individuals render to each other, through mercantile exchange, in view of their own interests, so that money becomes more important than virtue for the functioning of society (ibid., ii.349). This indicates the way in which Mandeville was founding a discourse on commerce on compromised moral principles - and, in that sense, inadequate as a source of legitimation of commercial society. It is interesting to note, however, how, in Smith’s view, Rousseau explores this aspect and elaborates, out of Mandevillian insights, a “sublime” - albeit exaggerated - critique of civilization (EPS, 251).

To justify the hypothesis of influence between Mandeville and Rousseau, Smith stresses some similarities between them, highlighting, however, important differences that lie at the basis of the theoretical move made by the latter. For both, men have no natural inclination to live in society, and despite their antagonistic descriptions of the state of nature, which lead them to think differently about the exit from this state, they both similarly describe the slow process by which men became sociable beings. Both think similarly about the origins of laws and, moreover, describe society as a space ruled by unsociable passions, such as the desire for superiority. However, Rousseau makes an important criticism of Mandeville regarding the conception of pity, in that he re-establishes the possibility of the moral virtues, whose existence had been rejected by the latter (see EPS, 250-1). According to Smith, Rousseau turned Mandeville’s principles into a critique of civilization through an antagonistic description of the state of nature, replete with rhetorical artifices, and a “philosophical chemistry”. The latter probably concerns the role he assigned to pity, describing it as one of the fundamental principles
of human action, alongside the impulse of self-preservation (amour de soi même), enabling the self-regulation of passions of natural men (Rousseau, 2002, pp. 84, 108). Thereby he could criticize Hobbes’ and Mandeville’s view of human nature, describing the natural state as a state of peace and isolation, and natural men as independent beings, with few needs and without an inclination to harm others.

As implied by Smith’s interpretation of the Second Discourse, Rousseau’s (2002, pp. 119-124) narrative suggests that men’s selfishness and unsociability are historical products, arising gradually with the development of human faculties and social cooperation. For him, pity, a feeling of aversion to the suffering of others, which drives natural beings to the aid of their fellows, is insufficient to unite individuals. Society is the result of accidents that drove them to coexistence oriented to the satisfaction of needs. The grouping of human beings gave rise to the emergence of language and cooperation, gradually awakening reason and comparison among individuals. Comparison, in turn, gave rise to self-consciousness through the perception of the opinions of others, which made them increasingly anxious to be seen and recognized by others. The awakening of self-love (amour-propre), understood essentially as the desire to be esteemed superior, gave rise to a growing spiral of psychological needs, which, with the advent of private property, led individuals to seek riches as a source of distinction. This was the root of the division of labor and social inequalities.

Therefore, for Rousseau, as suggested by Smith’s reading, what holds human beings together is not pity, but above all the desire for esteem and superiority, amour-propre, which arises together with social intercourse. The pursuit of distinction through wealth makes individuals dependent on the opinion, labor, and services of others, breaking the material and psychic independence of natural men, as well as places them in a state of competition and antagonism in society. As in Mandeville, dependence on others causes a split between being and appearing (ibid., p. 122), which makes social men false beings, willing to deceive others for their own benefit. Not by coincidence, these ideas appear in a relevant passage from the Second Discourse that was translated by Smith in his Letter (see EPS, 252-3).

Here, Rousseau critically describes the unsocial sociability of human “commerce”. Mutual social dependence causes individuals driven by amour-propre to

---

5 On this, see Force (2003, pp. 34-5) and Rasmussen (2006, pp. 632-3; 2008, pp. 66).
6 See also Rousseau (1999, pp. 78-9).
7 See the description of the process in Rousseau (2002, pp. 112-9).
interact by seeking to persuade others at any cost - “whether in reality or in appearance” - that it is in their interest to exchange on the terms proposed, implying a willingness to deceive others for their own benefit. This empties the moral content of exchange as a form of human interaction. Moreover, in Rousseau’s view, the division of labor, the advance of technology, the arts and sciences, and the production of wealth are achieved through inequality, oppression, and unhappiness. This “paradox” was already present, although without the character of a critique, in the well-known Mandevillian maxim. In the words of Colletti:

‘Private vices’ are ‘public benefits’ — this means not only that good is born of evil, that the immorality of individuals, their egoisms in competition with one another, produce culture and the ‘civilizing’ of society as a whole; it also means that wealth is born of poverty […], well-being from distress, that what produces prosperity is wage labour; or again that the wealth of a nation consists of a mass of toiling poor (Colletti, 1972, pp. 205).

Against Mandeville, “the most passionate detractor of human virtues” (Rousseau, 2002, pp. 106), Rousseau denounces the contradictory character of progress and criticizes, from a republican perspective, what Kant (2016 [1784], pp. 8-9) would later call men’s unsocial sociability (ungesellige Geselligkeit). As Smith points out in the Letter, Rousseau assumes Mandeville’s anthropology to show the morally and politically harmful effects arising from relations of social dependence founded on inequality and the sociability of amour-propre. Among them, as mentioned, are the threat to the integrity of the moral personality, the creation of oppressive inequalities, and the institution of an antagonism of interests that severs the political body.

3 The critique of moral egoism: pity, sympathy and sociability

In the Letter, Smith sets out to give a descriptive reading of Rousseau’s contribution. On the one hand, as seen, he recognizes some of the moves made by the latter in relation to Mandeville, such as the divergent description of the state of nature and the role ascribed to pity as the foundation of moral virtues, which allows him to criticize the Hobbesian-inspired conception of human nature. On the other hand, Rousseau would have accepted the Mandevillian description of the sociability of civilized individuals, using it as the basis of a “sublime” but excessive and partial critique of commerce (EPS,

8 On Rousseau’s denunciation of the contradictions between virtue and culture, personality and society, see Pocock (1975, pp. 502-5).
The partiality of this critique can be explained by Rousseau’s rhetorical style (EPS, 251). This, as noted Stimson (2015, pp. 356-8), is characterized by the attempt to persuade the reader by emphasizing the sides of the problem that endorse the argument one wants to establish and downplaying the sides that are contrary to it (LRBL, i.149-150). In this sense, in his description of the state of nature, Rousseau would not have adequately considered the negative aspects of the life of “savages”, and his critique of commerce and inequality would underestimate the benefits derived from the division of labor and civilization.

This reading helps us understand Smith’s move in TMS to think about the moral foundations of sociability out of a critical revision of the premises of Mandeville’s moral egoism. In Smith’s view, Rousseau was right to criticize the selfish anthropology and to recognize that it cannot found a good social order; and, nevertheless, he was wrong to incorporate it in his description of modern society. As we shall see, Smith tries to show precisely why Mandeville’s conception of the psychology behind the sociability of self-love is mostly wrong. But he does so not to deny the relevance of self-love, but to requalify it and rehabilitate in moral terms a discourse on society that was originally developed by authors close to the selfish tradition. Once this revision is made, an important part of Rousseau’s criticism of the contradictions of commercial societies becomes obsolete.

One might wonder why, despite many implicit references, Smith does not openly criticize Rousseau in the TMS, whereas he devotes an entire section of that work to Mandeville’s system. One possible reason is that it was enough to criticize Mandeville’s theses, since Rousseau’s arguments relied heavily on them. So that, by answering the former, he would have indirectly answered the latter. Another complementary reason is that, in his published works, Smith rarely explicitly criticizes contemporary thinkers that were still alive (probably as a matter of politeness). The most striking example is David Hume, who is probably the most important interlocutor of the whole book, but who is not mentioned once. All authors explicitly criticized by Smith in TMS were already dead.

---

9 According to Rasmussen (2018, pp. 253), this is what he means when he says that Rousseau’s republican spirit had been “carried a little too far” (EPS, 251). See also Winch (2002, pp. 301).

10 In our reading, this comment carries a critical tone, and therefore is not merely factual, as Force (2003, pp. 23-4) suggests.

11 In the essay Of the Imitative Arts (IA.24), Smith characterizes Rousseau as an author “more capable of feeling strongly than of analyzing accurately”.

12 There are still other instances that corroborate the latter reason. Smith also does not explicitly mention James Steuart in the Wealth of Nations, whose An Inquiry into the Principles of Political Oeconomy (1767) was one of his main targets (see CAS, 164; Smith, 1987). Another instance, which had been already noticed
Smith’s criticism of the doctrines of moral egoism, or the “systems which deduce the principle of approbation from self-love” (TMS VII.iii.1), among which those of Hobbes, Pufendorf, and Mandeville, are well known. Alongside Hutcheson and Hume, Smith argues that self-love is not the only motive of human action, nor the founding principle of moral distinctions; there are in human nature genuinely sociable affections and principles that make one to be disinterestedly concerned for the happiness of others (TMS I.i.1.1). Sympathy is one such principle, which enables one to put oneself in the other’s shoes, imagine what he or she feels and eventually feel with him or her, understand and evaluate their moral conduct in a disinterested way (TMS I.i.3.5, 10, I.i.3.1, VII.iii.1.4).13

In short, sympathy is an imaginative and reflexive capacity, by means of which one can project oneself into another’s perspective and circumstances and thereby form an idea of his feeling, being led (or not) to feel something similar, though to a lesser degree of intensity (TMS I.i.1.2). This is a non-selfish principle (TMS VII.iii.1.4), both in epistemic terms - it involves a genuine effort to imagine oneself being the other, and not merely a projection of oneself - and in moral terms - the ability to feel for others in a disinterested way, or to be genuinely concerned about the happiness of others (Griswold, 1999, p. 78-9). This capacity is at the core of the Smithian view of human nature and sociability, as well as of his theory of the origin of moral distinctions and social obligations.14

It is interesting to note how, in the first paragraphs of TMS, Smith arrives at sympathy starting from the concept of “pity or compassion, the emotion which we feel for the misery of others” (TMS I.i.1.1), in an implicit reference to the discussion made in the Letter15.

---

13 It is worth noting that moral judgment depends not only on the immediate identification of feelings provided by sympathy, but on conscious reflection and the use of reason (TMS III.4.8, VII.iii.2.6-7). On this, see Macfie (1967, pp. 64-5, 67), Griswold (1999, pp. 85, 88), and Montes (2004, pp. 47-8).

14 The concept of sympathy suffers some semantic variation in Smith’s text. Besides the meaning of a capacity that allows moral judgment, the term is sometimes used to designate the solidarity (fellow-feeling) of the spectator with the agent, or even a “correspondence of feelings” between both, produced from the imaginative exchange of perspectives (Cerqueira, 2008, p. 76). On the semantic of sympathy, see also Haakonsen (1981, p. 51) and Brown (2016). Montes (2004, p. 47-8) notes that it is a mistake to reduce sympathy to fellow-feeling, since it requires knowledge of the causes and circumstances of the action, involving reasoning and imagination.

15 This is also Stimson’s (2015, pp. 358-9) and Griswold’s (2018, pp. 130) opinion.
As some interpreters have noted\textsuperscript{16}, this indicates that, in a sense, Smith’s sympathy can be understood as a critical development of Rousseau’s concept of pity\textsuperscript{17}. In our reading, however, pity is only a starting point for Smith, being recognized as a feeling that attests to the existence of genuinely sociable principles in human nature. In itself, however, it is insufficient to found sociability, given that the latter is not primarily based on human misery, and that pity is a weak and short-lived feeling (LRBL ii.241), and does not constitute a motive capable of inspiring social agency. In this sense, Smith quickly draws attention to some distinctions between sympathy and pity, the first of which consists in the fact that sympathy is not restricted to feelings provoked by another’s suffering, but comprehends all passions (TMS I.i.1.4). In practice, sympathy contains pity, and Smith mobilizes the former to explain the latter.

Pity and compassion are words appropriated to signify our fellow-feeling with the sorrow of others. Sympathy, though its meaning was, perhaps, originally the same, may now, however, without much impropriety, be made use of to denote our fellow-feeling with any passion whatever (TMS I.i.1.5).

The second distinction is that sympathy does not depend on the direct sight of the display of the agent’s emotions\textsuperscript{18}. Although this can happen in some cases, in others, the direct perception of certain passions, without the knowledge of the circumstances which aroused them in the agent, may occasion antipathy and moral disapproval, as in the case of unsociable passions (TMS I.i.1.6-7). Smith’s thesis is that sympathy arises not so much from the sight of a passion in the agent as from knowledge of the circumstances that provoked it (TMS I.i.1.10).

As Griswold (2018, pp. 105-6) has noted, these distinctions highlight that sympathy is qualitatively different from pity. The former is a capacity grounded in imagination, which involves a degree of reflection and epistemic access to the feelings and motives of others, whereas the latter is a pre-reflective, pre-discursive, and pre-

\textsuperscript{16} Authors such as Winch (1996, pp. 72-3), Force (2003, pp. 19, 24, 28-9, 31-4), and Hont (2015, pp. 20-1, 33) stress the similarities between these concepts, suggesting that reading the \textit{Second Discourse} probably indicated to Smith that a generalization of pity was the way forward - although one should not overstate the hypothesis of influence, since Smith’s ideas were already relatively developed at that time (see Sagar, 2018a). Berry (2004), Hurtado (2005), Stimson (2015), and Griswold (2018, ch. 3), meanwhile, emphasize the differences between sympathy and pity, arguing that it makes more sense to think of the former as a critical response to the latter.

\textsuperscript{17} In our reading, it is not a matter of Rousseau’s influence on Smith, but of understanding what the Smithian conception represented in relation to that of Rousseau’s pity. The central influence, without any doubt, is Hume’s (cf. Rasmussen, 2017, pp. 90-4).

\textsuperscript{18} In \textit{The Treatise of Human Nature}, Hume (2009, pp. 404: T.2.2.7.4) states that pity depends on the direct sight of the object.
imaginative feeling\textsuperscript{19}. Sympathy is more complex than pity, not only because it encompasses all feelings, but because it involves imaginative operations that institute the possibility of a range of complex phenomena. Among them, the illusions of the mind upon the subject, which affect in a relevant way one’s behavior\textsuperscript{20}, and, as we shall see, morality itself.

It should be noted that these characteristics of pity, as represented in Rousseau’s \textit{Second Discourse}, make it a weak feeling, in a double sense. First, because it is insufficient to found sociability. In fact, on this point, Mandeville and Rousseau perfectly agree: for both, there is no innate principle of sociability in human nature; it is artificial, in the sense that it is acquired historically and involves unnatural passions\textsuperscript{21}. Secondly, pity is weak in the sense that it is overwhelmed and stifled in civilization by the competition for superiority and esteem, and is incapable of founding moral sentiments that can restrain the impetus of \textit{amour-propre}. Therefore, for Rousseau, spontaneous interaction driven by self-love is deceitful, devoid of genuine moral content, and potentially disruptive (Sagar, 2018b, pp. 158, 177).

As stated, Smith does not disagree with Rousseau that pity alone does not explain sociability and morality. Ultimately, sympathy supplants pity as a principle that characterizes human beings as disposed to sociability, and that does not wither but develops with civilization\textsuperscript{22}. As we shall see below, sympathy not only allows for the regulation of self-love, but gives it an ethical content, distinguishing it from Rousseau’s desire for superiority, and thereby making the search for recognition and competition in Smith of a non-disruptive character\textsuperscript{23}. Moreover, it is remarkable that Smith mobilizes this principle to explain (and legitimize, albeit in a qualified way)\textsuperscript{24} social stratification, since it disposes individuals to sympathize with others’ joy and prosperity, and thus to

\textsuperscript{19} Berry (2004, pp. 455-6) and Hurtado (2005, pp. 717) stress that the psychological mechanism of pity does not involve the imaginative exercise of projecting oneself into another’s perspective, body, and person that is characteristic of sympathy, so that the spectator remains in his or her own person.

\textsuperscript{20} On the illusive effects of the imagination upon the subject, see Griswold (1999, pp. 86-9) and Stimson (2015, pp. 364).

\textsuperscript{21} According to Griswold (2018, pp. 97-102, 107), Smith drew attention to the fact that Rousseau does not deny the Mandevillian idea of the selfish principle as the root of sociability, but merely historicizes it by removing it from natural beings.


\textsuperscript{23} On this, see Griswold (2018, pp. 124-5) and Sagar (2018b, pp. 176-8). Dupuy (1987, pp. 316-7, 339) and Force (2003, pp. 42, 165-8), on the other hand, bring Smith and Rousseau closer together on this point.

\textsuperscript{24} The legitimization of inequality, carried out in an effort to point out a source of stability in the social order, is not done without reservations, given that Smith adds that the disposition to sympathize with the rich and powerful is “the greatest and most universal cause of the corruption of our moral sentiments” (TMS Liii.3.1).
admire and respect the rich and powerful (TMS I.iii.2.1, 3) – in contrast to Rousseau’s critique of artificial inequalities among mankind.

While sympathy is a sociable, or at least non-selfish, principle, it should not be understood as a principle of action, opposed to self-love. Sympathy is qualitatively different from self-love: it is primarily a capacity allowed by the imagination, and not a particular feeling; in this sense, it is also distinct from benevolence, a sociable passion. Therefore, sympathy does not make human beings benevolent, but, as we shall see below, it enables them to regulate the natural unbalance of their feelings through the constitution of moral feelings. In its absence, human beings would be nothing but “wild beasts,” “and a man would enter an assembly of men as he enters a den of lions” (TMS II.ii.3.4). Be that as it may, the most important principle of action for Smith remains self-love, though a specific kind of it, namely the desire for moral recognition (cf. Bee, 2021).

In thinking of sympathy as a principle of sociability, Smith is not engaged with the question of the institution of society, like Mandeville and Rousseau. He simply assumes life in society as a given fact (Winch, 1996, p. 70), and on every occasion when he refers to some isolated individual living outside society, it is clearly a conjecture or thought experiment25. What is at stake is not how sociability arises historically, nor a conjectural history of society, but the process by which individuals become sociable as they interact spontaneously in society (that is, without the positive interference of government or the church). Smith shifts the discussion to explain how individuals naturally produce distinctions and moral rules in society from the operations of sympathy, independently and logically prior to any political or ecclesiastical authority.

4 Sympathy, impartial spectator and moral order

As a principle of sociability, sympathy constitutes the foundation of morality: it is through it that individuals become moral and self-conscious beings (Cerqueira, 2008, pp. 74-5). This occurs through two simultaneous and mirrored processes: i) individuals form moral distinctions and learn to judge from their own feelings as spectators of others’ conduct (TMS I.i.3.9); ii) they gradually become sensitive to the opinion others have of their conduct and begin to regulate their own passions so as to express them within the level appropriate to the approval of the spectators (TMS III.1.5; Griswold, 2018, pp. 105).

25 See TMS III.1.3, IV.1.8 and IV.2.12
In this process, the possibility of a moral self-assessment of our feelings and actions depends, in one way or another, on the perception of what are the feelings of others (TMS III.1.2). This is what is meant by Smith in the following thought experiment:

Were it possible that a human creature could grow up to manhood in some solitary place, without any communication with his own species, he could no more think of his own character, of the propriety or demerit of his own sentiments and conduct, of the beauty or deformity of his own mind, than of the beauty or deformity of his own face. All these are objects which he cannot easily see, which naturally he does not look at, and with regard to which he is provided with no mirror which can present them to his view. Bring him into society, and he is immediately provided with the mirror which he wanted before. It is placed in the countenance and behaviour of those he lives with, which always mark when they enter into, and when they disapprove of his sentiments; and it is here that he first views the propriety and impropriety of his own passions, the beauty and deformity of his own mind (TMS III.1.3).

This conjecture is intended to show that moral ideas, as well as self-consciousness, only take place through living with others. Society originally provides the individual with the “mirror” - the opinion and moral feelings of others - through which one becomes morally conscious of one’s own feelings and actions. To the extent that one becomes aware of the opinions others have about one’s conduct, one begins to wonder whether in fact one deserves their applause or censure. In order to evaluate oneself impartially, one must strive to examine one’s own conduct from a distance, imagining how it appears to others (TMS III.1.5). Moral conscience is formed through this process, which involves the formation of an imaginary internal mirror, through which the subject can see his own actions:

We suppose ourselves the spectators of our own behaviour, and endeavour to imagine what effect it would, in this light, produce upon us. This is the only looking-glass by which we can, in some measure, with the eyes of other people, scrutinize the propriety of our own conduct (TMS III.1.5).

To evaluate one’s own behavior, one performs a distancing from oneself and seeks to see oneself through the eyes of others (TMS III.1.1), something that is only possible through imagination. This mental effort founds self-reflexive sympathy, in which the agent observes and evaluates oneself as if from the perspective of another spectator. This engenders a kind of internal duplication between the person who acts and the person who judges one’s own conduct. The latter is a projection of the imagination that fulfills the function of an internal judge, potentially impartial by the fact that he judges from a certain distance from himself and has full knowledge of the circumstances that lead the subject (himself) to act. This is the concept of the “impartial spectator” (TMS III.1.6, III.2.30), which designates moral conscience, and which represents the way in which individuals
evaluate themselves by striving to view their own behavior as if they were external, disinterested spectators.

Gradually, by becoming self-conscious and observing the behavior of others, individuals form general rules regarding right and wrong, or what should or should not be done in particular situations (TMS III.4.7). They are based “upon experience of what, in particular instances, our moral faculties, our natural sense of merit and propriety, approve, or disapprove of” (TMS III.4.8), and are formed through inductive processes of reason (TMS VII.iii.2.6). Once formed in the mind, general rules can be used as criteria of what is just and unjust, and habitual respect for them instills in individuals a sense of duty, by which most of humanity regulates its conduct (TMS III.5.1). Thus, the rules founded on human feelings become moral laws themselves, sanctioned by conscience (TMS III.5.6), which delimit duties of justice and beneficence prior to any positive law or religious commandment.

Therefore, from the explanation of the formation of moral conscience mediated by sympathy, Smith arrives at the conception of an order governed by moral laws. As mentioned, the genesis of moral ideas, social obligations, and conscience is explained through the imaginary and mirrored “exchange” of feelings and perspectives of individuals who interact seeking recognition from others. From a political point of view, this means that individuals spontaneously produce the moral criteria relevant to the life in community, that is, without the tutelage of a sovereign or ecclesiastical power. This conception of a self-instituted order, in turn, means that moral conscience and social obligations are an unintended result of social interaction, and not something inscribed in human nature.

5 To be and to appear to be: Smith’s response to the problem of the moral status of self-love

Smith’s explanation of morality and conscience as socially constituted phenomena risked relativizing the source of moral criteria and social obligations. This is because he explained the distinction between right and wrong, just and unjust, as well as virtuous conduct, without recourse to an absolute moral criterion, but from the feelings and
opinions of mankind, and the desire for the esteem and sympathy of others\textsuperscript{26}. Now, Mandeville had done something similar by reducing human motivation to the desire for praise, or vanity (TMS VII.ii.4.7). In practice, Mandeville had reduced moral conscience to the desire to obtain esteem at any cost, eliminating the possibility of a genuinely virtuous motivation. This represented, in Rousseau’s terms, the enslavement of socialized individuals to the opinion of others. This thesis appears in one of the passages of the \textit{Second Discourse} translated by Smith in the \textit{Letter}: “For such in reality is the true cause of all those differences: the savage lives in himself; the man of society, always out of himself; cannot live but in the opinion of others, and it is, if I may say so, from their judgment alone that he derives the sentiment of his own existence” (EPS, 253).

Aware of this fact, Smith sought to bypass the idea, advocated by some “splenetic philosophers”, that moral motivation is reducible to the desire for unmerited esteem, or vanity (TMS III.2.27). Smith tried to show that, by becoming aware of others’ opinions about his own conduct, the moral subject begins to reflect on the extent to which he really deserves their applause or censure (TMS III.1.5). The judgment and regulation of one’s own conduct is mediated by the desire to be worthy of others’ approval, as distinct from the simple desire for esteem (TMS III.2.1-2)\textsuperscript{27}. This implies that the individual does not judge his feelings and actions primarily based on the opinion of real spectators, but has as his main reference what he admires in the actions of others, or what is worthy of the esteem of an impartial spectator. In this case, what motivates him is a love of self-approval, or a love of virtue (TMS III.2.8), which instills in human beings the desire \textit{to be} - and not just \textit{to appear to be} - sociable.

Nature, when she formed man for society, endowed him with an original desire to please, and an original aversion to offend his brethren. She taught him to feel pleasure in their favourable, and pain in their unfavourable regard. She rendered their approbation most flattering and most agreeable to him for its own sake; and their disapprobation most mortifying and most offensive. But this desire of the approbation, and this aversion to the disapprobation of his brethren, would not alone have rendered him fit for that society for which he was made. Nature, accordingly, has endowed him, not only with a desire of being approved of, but with a desire of being what ought to be approved of; or of being what he himself approves of in other men. The first desire could only have made him wish \textit{to appear to be fit for society}. The second was necessary in order to render him anxious \textit{to be really fit} (TMS III.2.6-7, our emphasis).

\textsuperscript{26} On this, see Hont (2015, pp. 29–32), who presents some of the most important criticisms of Smith’s moral theory made by Adam Ferguson, Thomas Reid, and Dugald Stewart.

\textsuperscript{27} It is important to note that, according to Raphael (1975) and Hanley (2009), Smith developed this aspect of his theory of the impartial spectator throughout his life, and made substantive changes to the editions of TMS, particularly the second and sixth editions.
We see here, again, an implicit dialogue with Mandeville and Rousseau. As seen, the Mandevillian schism between being and appearing implied, in Rousseau’s terms, the corruption of the human being\(^{28}\), the reduction of sociability to an instrument of egoism. As if individuals lived together only because of the material advantages they obtained and the pleasure of seeing themselves esteemed superior at any cost. Individuals would be essentially selfish and only apparently sociable, regulating their own conduct not according to what is right and just as ends in themselves, but only for the pleasure of obtaining esteem and material benefits from others. As is clear from the passage above, in Smith’s view, the desire for esteem alone is incapable of making human beings fit for life in society.

The theory of the impartial spectator and the distinction between the desire for esteem and the desire to be praiseworthy can be seen as an attempt to answer the question of the moral status of sociability\(^{29}\). Smith criticizes Mandeville precisely for the fact that he reduces moral motivation to vanity, or the desire for undeserved esteem (TMS VII.ii.4.7). In contrast, he distinguishes vanity from the desire for deserved esteem - the desire for true glory, which can also be called the desire for moral recognition - and from the desire to be worthy of others’ esteem, also referred to as the desire for self-approval, or love of virtue, which does not involve the desire to actually obtain public esteem (TMS VII.ii.4.8-10). As Bee (2021) argues, for Smith, the main driver of moral behavior and sociability is the desire for moral recognition, that is, the desire to obtain the esteem of others on the basis of meritorious qualities\(^{30}\). This is essentially the desire to better one’s condition (TMS I.iii.2.1, WN II.iii.28), as well as the moral motive behind exchange (Bee, 2021, pp. 124-131). That is, for Smith, exchange-based commercial interaction is not founded on the desire to persuade others at any cost, but on the desire to obtain recognition for one’s true merit. To receive an equivalent in exchange for one’s services means that the other recognizes the merit of one’s good offices, and therefore confirms one’s self-esteem.

\(^{28}\) See Hanley (2009, pp. 30-1, 41-2).

\(^{29}\) As noted by Stimson (2015, p. 361, 364), this is tied to the question posed by Mandeville about the absence of a normative distance between one’s judgment of his own conduct and the opinion of others. On this, see also Hundert (1994, pp. 227) and Griswold (2018, pp. 123).

\(^{30}\) According to Smith, the desire for unmerited esteem is restricted to a few vain individuals, or to a few moments in one’s moral experience, whereas the love of virtue, that is, the desire to be worthy of the approval of an impartial spectator without the need to obtain public esteem, is restricted to a few wise men (TMS III.2.7, 11, 28; Bee, 2021, pp. 122-3).
Likewise, Smith emphasizes that the pleasure derived from moral recognition, or from the perception that spectators sympathize with our meritorious qualities, is immediate, or pre-reflexive (TMS I.i.2.1). This means that the pleasure elicited by the experience of mutual recognition is not derived from an interested calculation - as if it was a result from the subject’s reasoning that the sympathy of others will bring him benefits in the future. Hence, the pleasure associated with sociability is not selfish or instrumental, provided that by selfishness is meant the search for advantages derived from things that are useful to us, without regard for the moral feelings of others31. In other words, sociability and the desire for recognition do not have as their purpose a simple utility or material advantage.

Therefore, Mandeville is wrong to assume that individuals want to appear to be what they are not in order to satisfy their selfish desires and interests. In Smith’s view, most seek to be what they appear to be, or appear to be what they really are, that is, they seek to obtain the esteem of others based on qualities they really possess. Moral subjects, although dependent on the opinion of actual spectators, are not enslaved to it. That is, they do not submit to it at any cost, but can preserve the authenticity of their personality and the independence of their judgment to the extent that they adopt as a criterion the opinion of impartial and disinterested spectators.

According to Griswold (2018, pp. 125, 128), this means that although self-consciousness depends on consideration for the opinion of others, this does not impugn one’s moral authenticity, but is its condition of possibility. So that dependence on the opinion of others in general does not translate itself into the constitution of a false sociability, characterized by deceit, manipulation, and hypocrisy (as in Mandeville and Rousseau), or by isolation and an unsociable competition for status. On the contrary, it allows the sociability of self-love to acquire a truly ethical content, in which individuals seek in the opinion of others the confirmation of their own judgment as to their worth and the value of what they produce.

6 Conclusion

31 This is the definition of a natural good, according to Hutcheson (2004, pp. 85-6), in distinction to a moral good, understood as the pleasure derived from the observation of actions endowed with moral value. Cf. Cerqueira (2008, pp. 65). Smith sometimes seems to operate implicitly with this distinction (cf. TMS I.iii.2.1, IV.2.12).
Rousseau’s critique of the selfish anthropology is relevant for Smith, insofar as it illustrates the harmful consequences of the Mandevillian idea of society. At the same time, however, according to Smith, Rousseau’s critique of commerce and civilization is largely unfounded because it is based on a mostly mistaken moral psychology. In Smith’s view, Rousseau would be right if Mandeville’s theses on sociability and self-love were correct. In this case, human “commerce” would not have a genuine moral content, since the competition for status and superiority driven by the desire for approval would dispose individuals to deceive others for their own benefit. Interaction would be the realm of false appearances and veiled opposition of interests, being devoid of an ethical content of its own.

For Smith, on the contrary, sociability acquires a genuinely moral character insofar as the object of self-love is the deserved recognition of others. Smith’s theory of sympathy and the impartial spectator is a response to Mandeville’s idea of the desire for praise as the vehicle of sociability, as it is insufficient to make human beings fit for life in society. In Smith’s view, there is no split between essence and appearance in the process of socialization: human beings actually become sociable as they seek to be that which is worthy of the esteem of an impartial spectator. Society and human “commerce” are thought of as an order governed by moral laws, arising from sociability itself and thus having a relative autonomy from political and ecclesiastical authority.

Starting from this conception of sociability, it is possible for Smith to think about the objects of political economy – such as commerce, exchange, production, distribution and accumulation of wealth – as phenomena governed by rules of their own. That is, as objects endowed with a legality relatively independent from the impositions of a superior power, amenable to be theorised on their own. This, however, does not mean that, for Smith, these phenomena are separated from the moral dimension of social life. On the contrary, they can achieve a certain autonomy from the political only to the extent they are considered a form of manifestation of genuinely moral interaction.

References


About the author

Ivan Prates Sternick – ivansternick@gmail.com

Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil.

ORCID: https://orcid.org/0000-0002-7957-0174

Acknowledgments

This paper is a result of my dissertation research, and for this I thank my master’s advisor, Prof. Helton Adverse, and the members of the panel, Professors Hugo Cerqueira and Newton Bignotto. I also thank the colleagues of the research group on the history of economic thought at Cedeplar/UFMG, and especially Professor Michele Bee for the discussion of the text. Finally, I thank the editors and referees of Nova Economia. The opinions here expressed are of my exclusive responsibility. This work was supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).
About the article
Submission received on April 06, 2022. Approved for publication on June 27, 2022.

Conflict of interest statement
The authors declare that there are no conflicts of interest of personal, commercial, academic or financial nature between the authors and the publication of this article.
This preprint was submitted under the following conditions:

- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.

- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.

- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.

- The authors declare that the data, applications, and other content underlying the manuscript are referenced.

- The deposited manuscript is in PDF format.

- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.

- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.

- The authors agree that the approved manuscript will be made available under a [Creative Commons CC-BY](https://creativecommons.org/licenses/by/4.0/) license.

- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.

- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.

- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.

- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.